

PCT09

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/806,876A

DATE: 10/23/2001
 TIME: 10:15:00

Input Set : A:\ES.txt
 Output Set: N:\CRF3\10232001\I806876A.raw

4 <110> APPLICANT: Ress-Loeschke, Marion
 5 Friedrich, Thomas
 6 Hauer, Bernhard
 8 <120> TITLE OF INVENTION: A process for preparing chiral carboxylic acids
 9 from nitriles using a nitrilase or microorganisms
 10 which comprise a gene for the nitrilase
 12 <130> FILE REFERENCE: 49462
 W--> 13 <140> CURRENT APPLICATION NUMBER: US 09/806,876A
 14 <141> CURRENT FILING DATE: 2001-05-21
 15 <150> PRIOR APPLICATION NUMBER: Germany/19848129.2
 16 <151> PRIOR FILING DATE: 1998-10-19
 17 <160> NUMBER OF SEQ ID NOS: 9
 18 <170> SOFTWARE: WordPerfect version 6.1
 19 <210> SEQ ID NO: 1
 20 <211> LENGTH: 1071
 21 <212> TYPE: DNA
 22 <213> ORGANISM: *Alcaligenes faecalis*
 23 <220> FEATURE:
 24 <221> NAME/KEY: CDS
 25 <222> LOCATION: 1 ... 1071
 26 <400> SEQUENCE: 1
 14 atg cag aca aga aaa atc gtc cgg gca gcc gcc gta cag gcc gcc tct 48
 15 Met Gln Thr Arg Lys Ile Val Arg Ala Ala Ala Val Gln Ala Ala Ser
 16 1 5 10 15
 36 1 5 10 15 96
 38 ccc aac tac gat ctg gca acg ggt gtt gat aaa acc att gag ctg gct
 39 Pro Asn Tyr Asp Leu Ala Thr Gly Val Asp Lys Thr Ile Glu Leu Ala
 40 20 25 30
 42 cgt cag gcc cgc gat gag ggc tgt gac ctg atc gtg ttt ggt gaa acc 144
 43 Arg Gln Ala Arg Asp Glu Gly Cys Asp Leu Ile Val Phe Gly Glu Thr
 44 35 40 45
 46 tgg ctg ccc gga tat ccc ttc cac gtc tgg ctg ggc gca ccg gcc tgg 192
 47 Trp Leu Pro Gly Tyr Pro Phe His Val Trp Leu Gly Ala Pro Ala Trp
 48 50 55 60
 50 tcg ctg aaa tac agt gcc cgc tac tat gcc aac tcg ctc tcg ctg gac 240
 51 Ser Leu Lys Tyr Ser Ala Arg Tyr Tyr Ala Asn Ser Leu Ser Leu Asp
 52 65 70 75 80
 54 agt gca gag ttt caa cgc att gcc cag gcc gca cgg acc ttg ggt att 288
 55 Ser Ala Glu Phe Gln Arg Ile Ala Gln Ala Ala Arg Thr Leu Gly Ile
 56 85 90 95
 58 ttc atc gca ctg ggt tat agc gag cgc agc ggc ggc agc ctt tac ctg 336
 59 Phe Ile Ala Leu Gly Tyr Ser Glu Arg Ser Gly Gly Ser Leu Tyr Leu
 60 100 105 110
 62 ggc caa tgc ctg atc gac gac aag ggc gag atg ctg tgg tcg cgt cgc 384
 63 Gly Gln Cys Leu Ile Asp Asp Lys Gly Glu Met Leu Trp Ser Arg Arg
 64 115 120 125
 66 aaa ctc aaa ccc acg cat gta gag cgc acc gta ttt ggt gaa ggt tat 432
 67 Lys Leu Lys Pro Thr His Val Glu Arg Thr Val Phe Gly Glu Gly Tyr

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/806,876A

DATE: 10/23/2001

TIME: 10:15:00

Input Set : A:\ES.txt

Output Set: N:\CRF3\10232001\I806876A.raw

```

68      130      135      140
70 gcc cgt gat ctg att gtg tcc gac aca gaa ctg gga cgc gtc ggt gct      480
71 Ala Arg Asp Leu Ile Val Ser Asp Thr Glu Leu Gly Arg Val Gly Ala
72 145      150      155      160
74 cta tgc tgc tgg gag cat ttg tgc ccc ttg agc aag tac gcg ctg tac      528
75 Leu Cys Cys Trp Glu His Leu Ser Pro Leu Ser Lys Tyr Ala Leu Tyr
76      165      170      175
78 tcc cag cat gaa gcc att cac att gct gcc tgg ccg tgc ttt tgc cta      576
79 Ser Gln His Glu Ala Ile His Ile Ala Ala Trp Pro Ser Phe Ser Leu
80      180      185      190
82 tac agc gaa cag gcc cac gcc ctc agt gcc aag gtg aac atg gct gcc      624
83 Tyr Ser Glu Gln Ala His Ala Leu Ser Ala Lys Val Asn Met Ala Ala
84      195      200      205
86 tgc caa atc tat tgc gtt gaa ggc cag tgc ttt acc atc gcc gcc agc      672
87 Ser Gln Ile Tyr Ser Val Glu Gly Gln Cys Phe Thr Ile Ala Ala Ser
88      210      215      220
90 agt gtg gtc acc caa gag acg cta gac atg ctg gaa gtg ggt gaa cac      720
91 Ser Val Val Thr Gln Glu Thr Leu Asp Met Leu Glu Val Gly Glu His
92 225      230      235      240
94 aac gcc ccc ttg ctg aaa gtg ggc ggc ggc agt tcc atg att ttt gcg      768
95 Asn Ala Pro Leu Leu Lys Val Gly Gly Gly Ser Ser Met Ile Phe Ala
96      245      250      255
98 ccg gac gga cgc aca ctg gct ccc tac ctg cct cac gat gcc gag ggc      816
99 Pro Asp Gly Arg Thr Leu Ala Pro Tyr Leu Pro His Asp Ala Glu Gly
100      260      265      270
102 ttg atc att gcc gat ctg aat atg gag gag att gcc ttc gcc aaa gcg      864
103 Leu Ile Ile Ala Asp Leu Asn Met Glu Glu Ile Ala Phe Ala Lys Ala
104      275      280      285
106 atc aat gac ccc gta ggc cac tat tcc aaa ccc gag gcc acc cgt ctg      912
107 Ile Asn Asp Pro Val Gly His Tyr Ser Lys Pro Glu Ala Thr Arg Leu
108      290      295      300
110 gtg ctg gac ttg ggg cac cga gac ccc atg act cgg gtg cac tcc aaa      960
111 Val Leu Asp Leu Gly His Arg Asp Pro Met Thr Arg Val His Ser Lys
112 305      310      315      320
114 agc gtg acc agg gaa gag gct ccc gag caa ggt gtg caa agc aag att      1008
115 Ser Val Thr Arg Glu Glu Ala Pro Glu Gln Gly Val Gln Ser Lys Ile
116      325      330      335
118 gcc tca gtc gct atc agc cat cca cag gac tgc gac aca ctg cta gtg      1056
119 Ala Ser Val Ala Ile Ser His Pro Gln Asp Ser Asp Thr Leu Leu Val
120      340      345      350
122 caa gag ccg tct tga      1071
123 Gln Glu Pro Ser
124      355
127 <210> SEQ ID NO: 2
128 <211> LENGTH: 356
129 <212> TYPE: PRT
130 <213> ORGANISM: Alcaligenes faecalis
132 <400> SEQUENCE: 2
134 Met Gln Thr Arg Lys Ile Val Arg Ala Ala Ala Val Gln Ala Ala Ser

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/806,876A

DATE: 10/23/2001

TIME: 10:15:00

Input Set : A:\ES.txt

Output Set: N:\CRF3\10232001\I806876A.raw

```

135      1              5              10              15
137 Pro Asn Tyr Asp Leu Ala Thr Gly Val Asp Lys Thr Ile Glu Leu Ala
138              20              25              30
140 Arg Gln Ala Arg Asp Glu Gly Cys Asp Leu Ile Val Phe Gly Glu Thr
141              35              40              45
143 Trp Leu Pro Gly Tyr Pro Phe His Val Trp Leu Gly Ala Pro Ala Trp
144              50              55              60
146 Ser Leu Lys Tyr Ser Ala Arg Tyr Tyr Ala Asn Ser Leu Ser Leu Asp
147 65              70              75              80
149 Ser Ala Glu Phe Gln Arg Ile Ala Gln Ala Ala Arg Thr Leu Gly Ile
150              85              90              95
152 Phe Ile Ala Leu Gly Tyr Ser Glu Arg Ser Gly Gly Ser Leu Tyr Leu
153 100              105              110
155 Gly Gln Cys Leu Ile Asp Asp Lys Gly Glu Met Leu Trp Ser Arg Arg
156 115              120              125
158 Lys Leu Lys Pro Thr His Val Glu Arg Thr Val Phe Gly Glu Gly Tyr
159 130              135              140
161 Ala Arg Asp Leu Ile Val Ser Asp Thr Glu Leu Gly Arg Val Gly Ala
162 145              150              155              160
164 Leu Cys Cys Trp Glu His Leu Ser Pro Leu Ser Lys Tyr Ala Leu Tyr
165 165              170              175
167 Ser Gln His Glu Ala Ile His Ile Ala Ala Trp Pro Ser Phe Ser Leu
168 180              185              190
170 Tyr Ser Glu Gln Ala His Ala Leu Ser Ala Lys Val Asn Met Ala Ala
171 195              200              205
173 Ser Gln Ile Tyr Ser Val Glu Gly Gln Cys Phe Thr Ile Ala Ala Ser
174 210              215              220
176 Ser Val Val Thr Gln Glu Thr Leu Asp Met Leu Glu Val Gly Glu His
177 225              230              235              240
179 Asn Ala Pro Leu Leu Lys Val Gly Gly Gly Ser Ser Met Ile Phe Ala
180 245              250              255
182 Pro Asp Gly Arg Thr Leu Ala Pro Tyr Leu Pro His Asp Ala Glu Gly
183 260              265              270
185 Leu Ile Ile Ala Asp Leu Asn Met Glu Glu Ile Ala Phe Ala Lys Ala
186 275              280              285
188 Ile Asn Asp Pro Val Gly His Tyr Ser Lys Pro Glu Ala Thr Arg Leu
189 290              295              300
191 Val Leu Asp Leu Gly His Arg Asp Pro Met Thr Arg Val His Ser Lys
192 305              310              315              320
194 Ser Val Thr Arg Glu Glu Ala Pro Glu Gln Gly Val Gln Ser Lys Ile
195 325              330              335
197 Ala Ser Val Ala Ile Ser His Pro Gln Asp Ser Asp Thr Leu Leu Val
198 340              345              350
200 Gln Glu Pro Ser
201 355
204 <210> SEQ ID NO: 3
205 <211> LENGTH: 39
206 <212> TYPE: PRT
207 <213> ORGANISM: Alcaligenes faecalis

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/806,876A

DATE: 10/23/2001
TIME: 10:15:00

Input Set : A:\ES.txt
Output Set: N:\CRF3\10232001\I806876A.raw

```

209 <400> SEQUENCE: 3
211 Met Gln Thr Arg Lys Ile Val Arg Ala Ala Ala Val Gln Ala Ala Ser
212   1             5             10             15
214 Pro Asn Tyr Asp Leu Ala Thr Gly Val Asp Lys Thr Ile Glu Leu Ala
215             20             25             30
217 Arg Gln Ala Arg Asp Glu Gly
218             35
221 <210> SEQ ID NO: 4
222 <211> LENGTH: 21
223 <212> TYPE: PRT
224 <213> ORGANISM: Alcaligenes faecalis
226 <400> SEQUENCE: 4
228 Glu Glu Ala Pro Glu Gln Gly Val Gln Ser Lys Ile Ala Ser Val Ala
229   1             5             10             15
231 Ile Ser His Pro Gln
232             20
235 <210> SEQ ID NO: 5
236 <211> LENGTH: 11
237 <212> TYPE: PRT
238 <213> ORGANISM: Alcaligenes faecalis
240 <400> SEQUENCE: 5
242 Glu Glu Ala Pro Glu Gln Gly Val Gln Ser Lys
243   1             5             10
246 <210> SEQ ID NO: 6
247 <211> LENGTH: 23
248 <212> TYPE: DNA
249 <213> ORGANISM: Alcaligenes faecalis
251 <220> FEATURE:
252 <221> NAME/KEY: Unsure
253 <222> LOCATION: 1 ... 23 OK
254 <223> OTHER INFORMATION: n represents g, a, t or c
256 <400> SEQUENCE: 6
W--> 259 atgcagacna gnaaratcgt scg 23
262 <210> SEQ ID NO: 7
263 <211> LENGTH: 20
264 <212> TYPE: DNA
265 <213> ORGANISM: Alcaligenes faecalis
267 <220> FEATURE:
268 <221> NAME/KEY: Unsure
269 <222> LOCATION: 1 ... 20 OK
270 <223> OTHER INFORMATION: n represents g, a, t or c
273 <400> SEQUENCE: 7
W--> 275 tngcsacnga ngcratcttg 20
278 <210> SEQ ID NO: 8
279 <211> LENGTH: 31
280 <212> TYPE: DNA
281 <213> ORGANISM: Alcaligenes faecalis
283 <400> SEQUENCE: 8
285 ttaatcatat gcagacaaga aaaatcggtcc g 31

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/806,876A

DATE: 10/23/2001
TIME: 10:15:00

Input Set : A:\ES.txt
Output Set: N:\CRF3\10232001\I806876A.raw

288 <210> SEQ ID NO: 9
289 <211> LENGTH: 32
290 <212> TYPE: DNA
291 <213> ORGANISM: Alcaligenes faecalis
293 <400> SEQUENCE: 9
295 aaggatcctc aagacggctc ttgcactagc ag

32

T05040-92890860

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/806,876A

DATE: 10/23/2001

TIME: 10:15:01

Input Set : A:\ES.txt

Output Set: N:\CRF3\10232001\I806876A.raw

L:13 M:283 W: Missing Blank Line separator, <140> field identifier

L:259 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6

L:275 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7

FO50110-92890350